Fringe Field Effects and Optimal Symplectification of Maps

Béla Erdélyi

Department of Physics and Astronomy,
National Superconducting Cyclotron Laboratory,
Michigan State University,
East Lansing, MI 48824, USA

LBNL, CBP, June 1, 2001

Motivation

To develop (long term) tracking methods that are:

- Accurate
- Efficient

Symplectic tracking with one-turn maps:

1. Compute truncated one-turn Taylor map
2. Symplectify it
3. Iterate the resulting exactly symplectic map

It is very important to incorporate all relevant effects in the one-turn map, and to symplectify it in the right way

$\Downarrow$

Accurate Fringe Field Effects and
Optimal Symplectification

On the fringe ...

Fringe field effects are often neglected or simplistic models are employed

The fringe fields are sometimes important:

- Noticeable, but not a limiting factor (LHC)
- The effects tend to become especially significant for small rings and large emittances (Muon Collider and Neutrino Factory)

Generic fringe field effects:

- aperture dependent
- fall-off shape dependent
- sharp cutoff leads to divergences

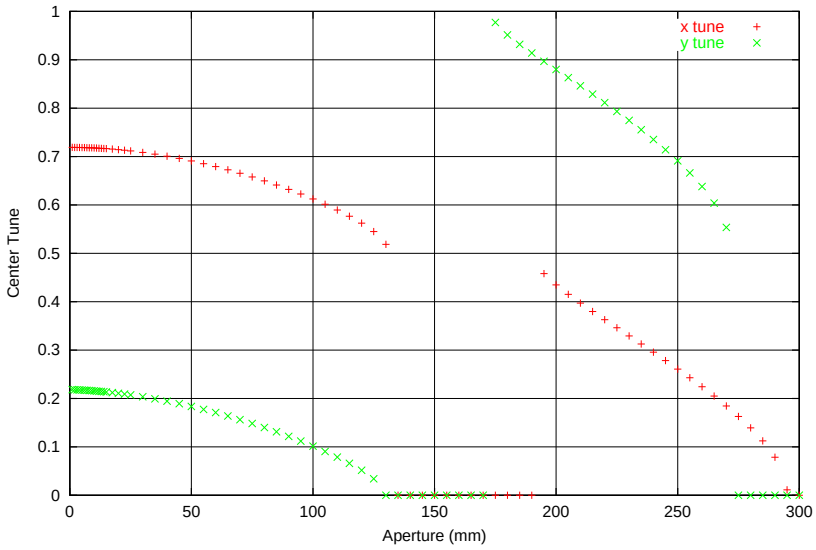

Figure 1: Center tunes as a function of aperture.

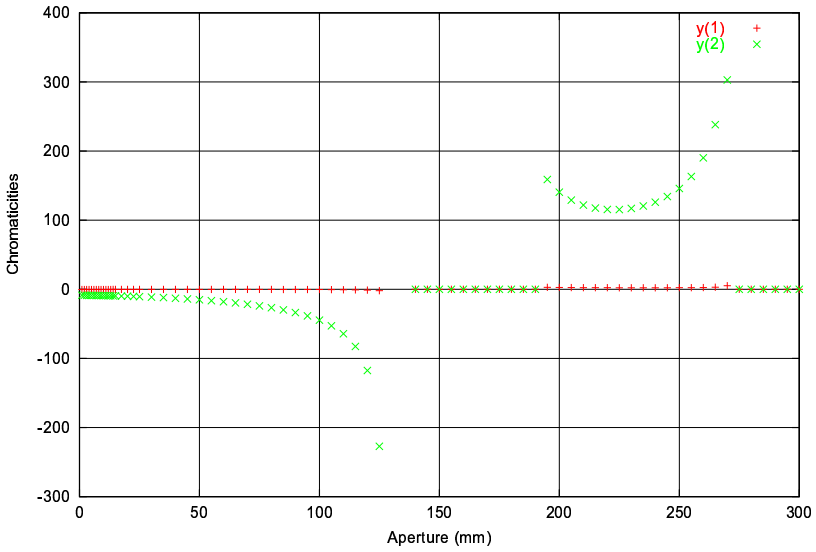

Figure 2: First-order and second-order y-chromaticities as a function of aperture.

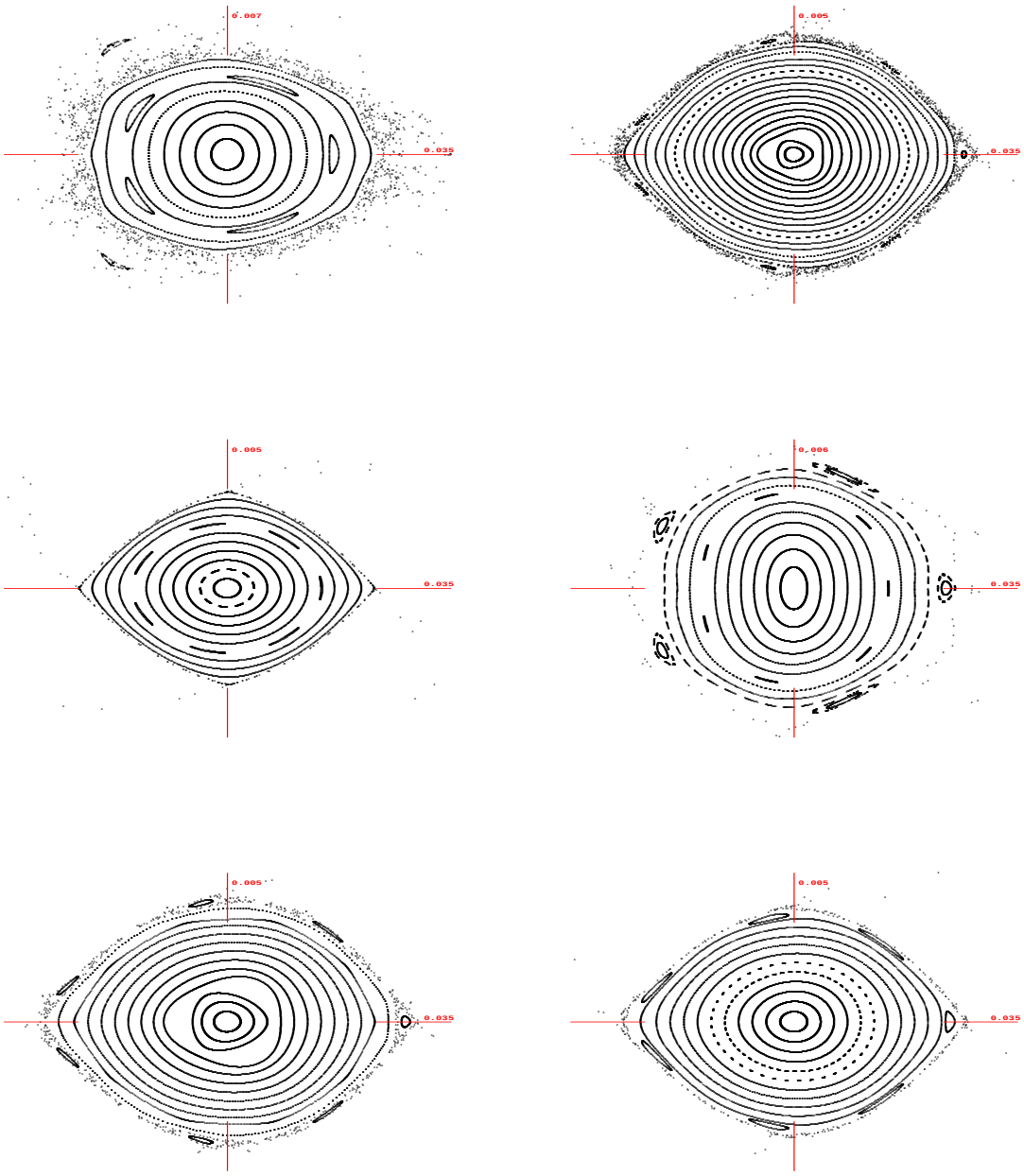

Figure 3: Tracking pictures for six different fringe field shapes.

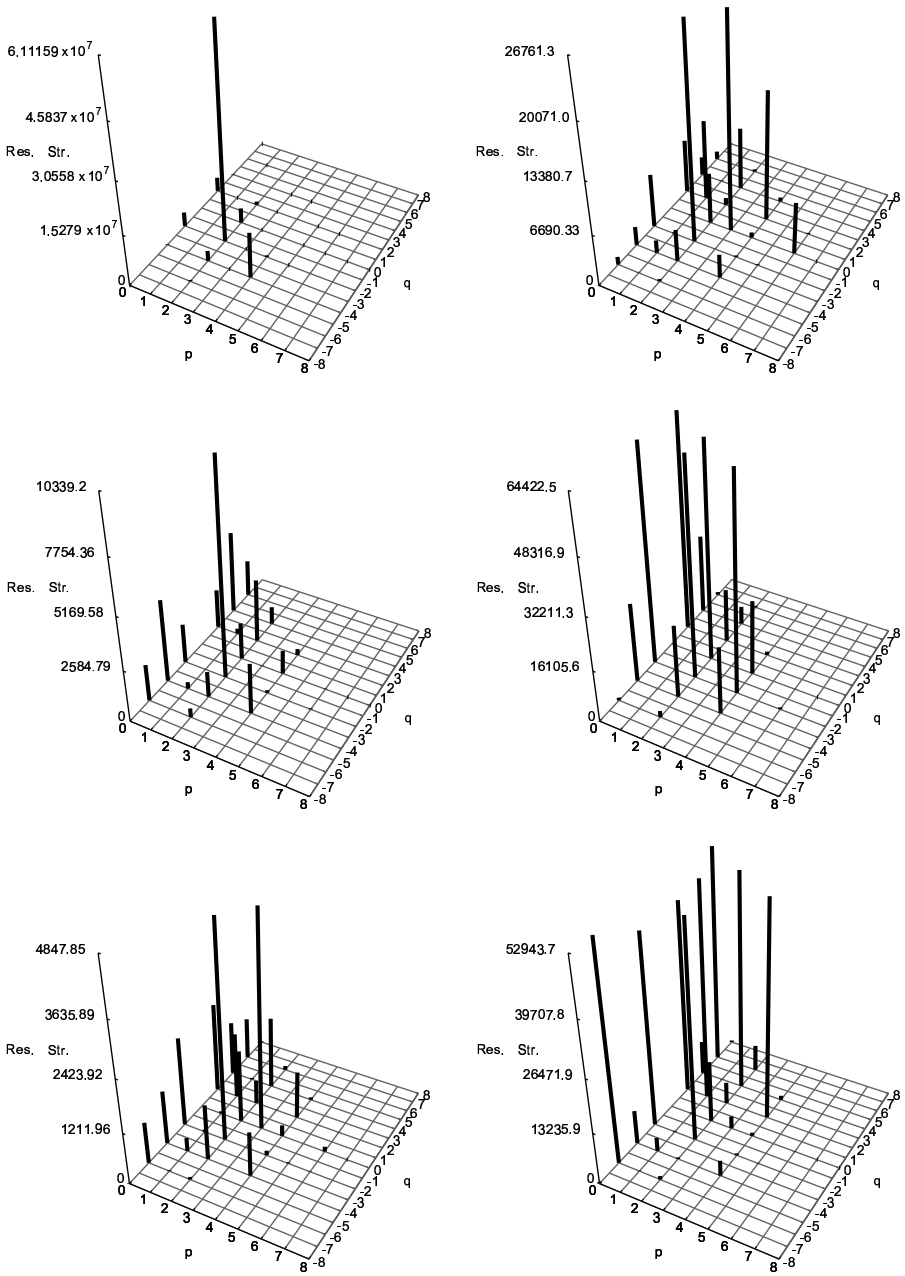

Figure 4: Resonance strengths for the same six fringe field shapes.

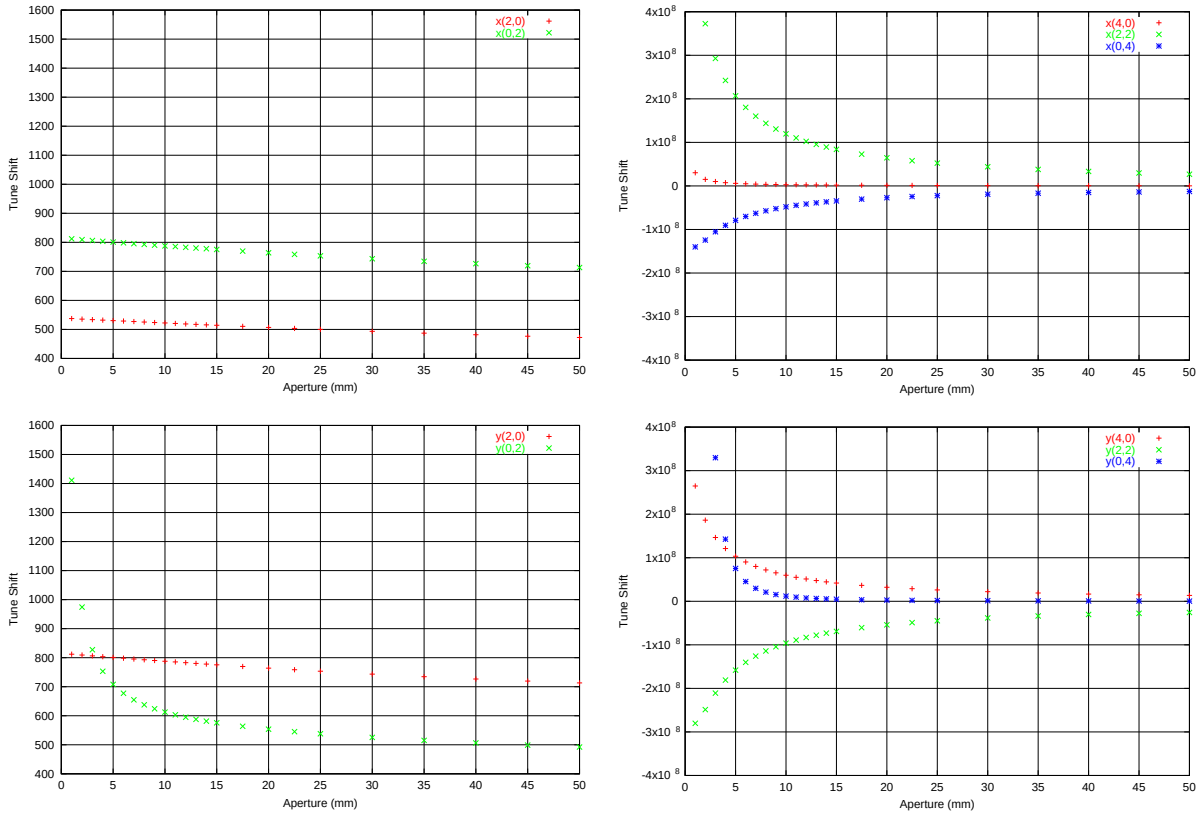

Figure 5: Blowup of the amplitude dependent tuneshifts in the sharp cutoff approximation.

Application to the Neutrino Factory and the Proton Driver

- Generic fringe field shape
- More detrimental for the Neutrino Factory than the Proton Driver

Conclusions

- Incorrect treatment of fringe and other nonlinear effects can lead to incorrect prediction of dynamic behavior
- It is wise to study fringe field effects on a case-by-case basis

$\Rightarrow$ **Need method for "exact" fringe field map computation**

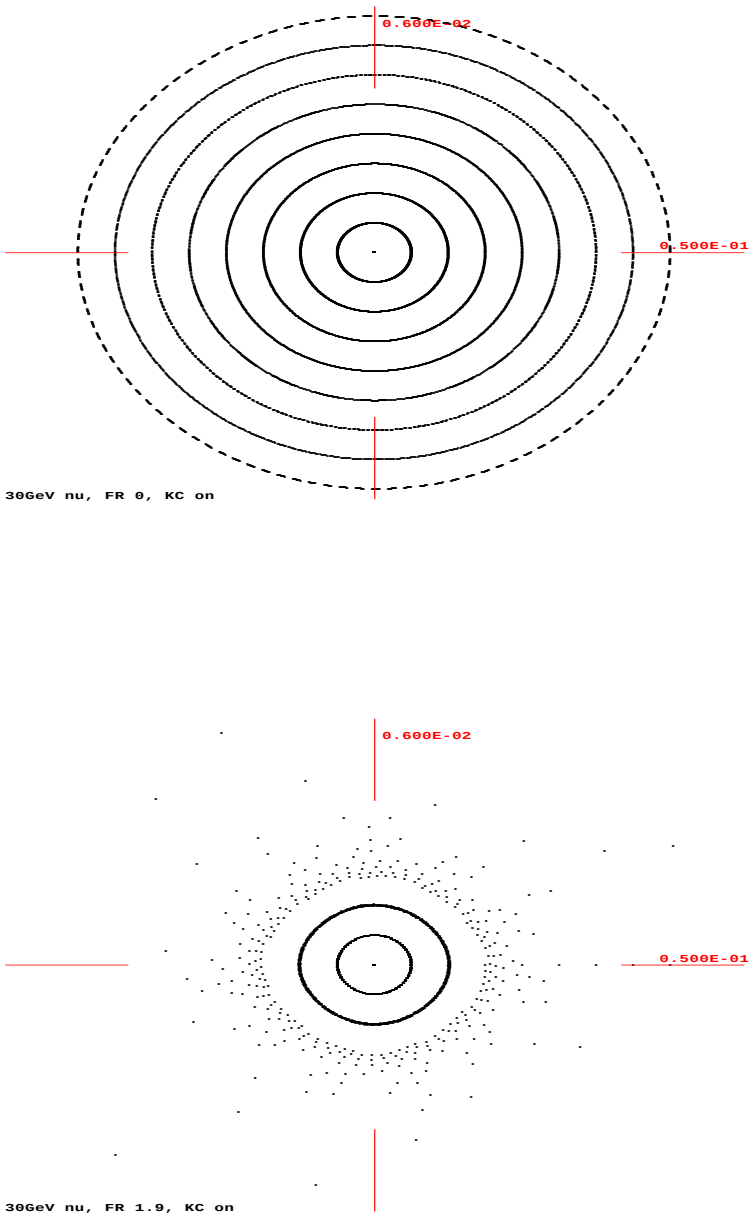

Figure 6: Tracking of the Neutrino Factory without and with fringe field effects.

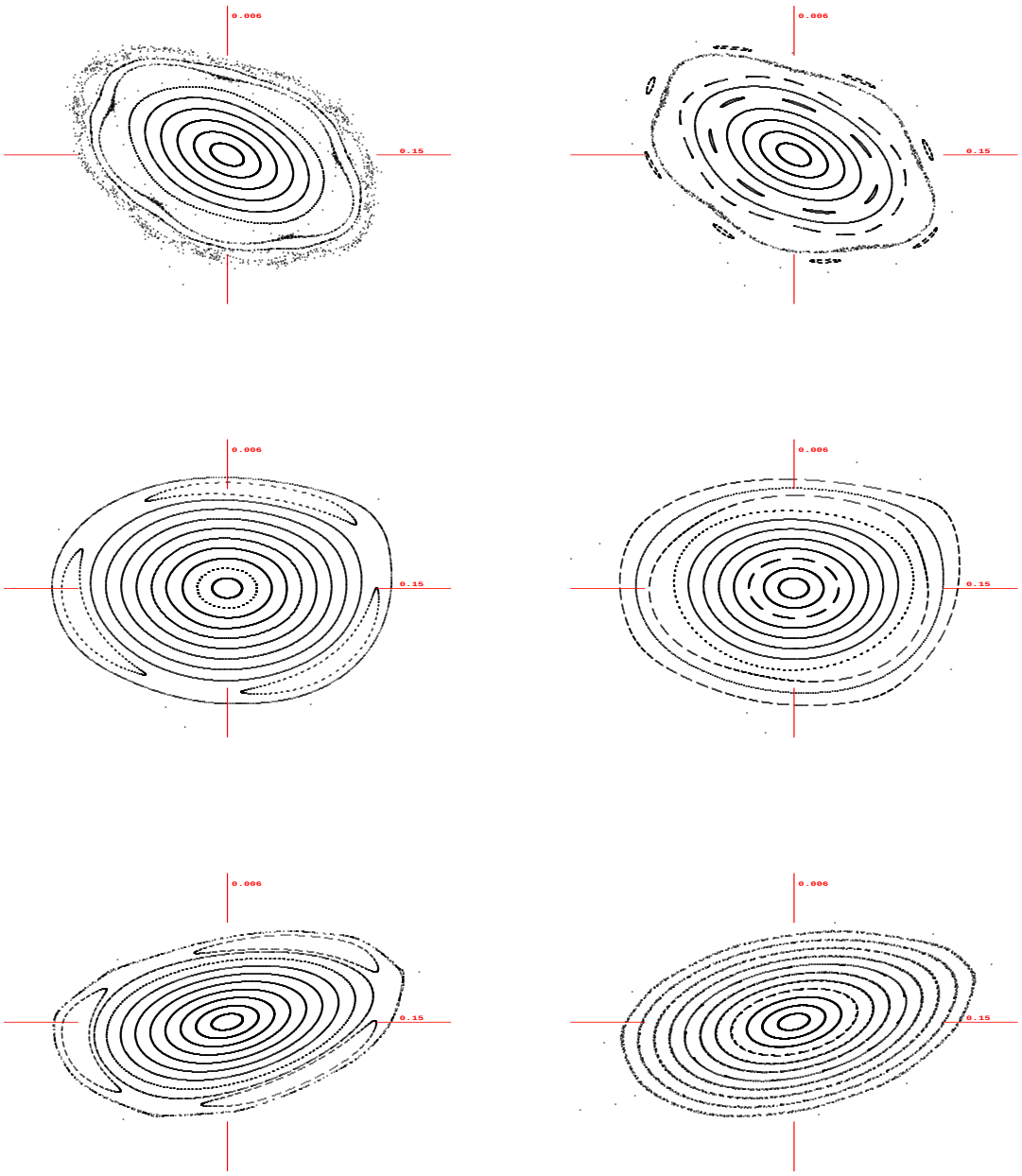

Figure 7: Tracking pictures of the Proton Driver without and with fringe fields, and momentum offsets of 0 and $\pm 4\%$.

Differential Algebra-based multipole decomposition

Applicable to current dominated superconducting magnets, for which analytical magnet models exist

Properties:

- It is exact within the model
- It is the only practical analytic method that works to very high orders
- It is easy to check and enforce Maxwell's equations

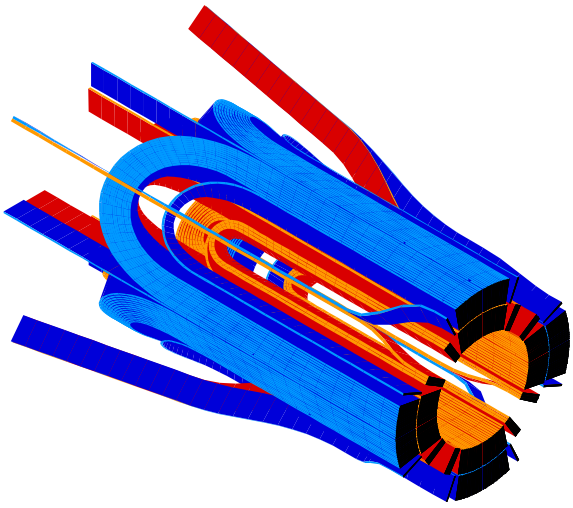

Figure 8: Lead End of the High Gradient Quadrupoles of the Large Hadron Collider.

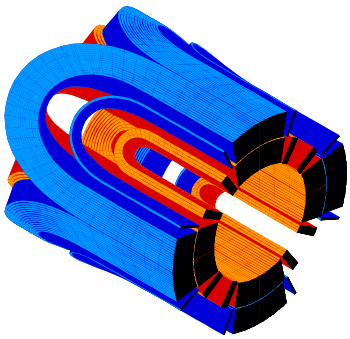

Figure 9: Return End of the High Gradient Quadrupoles of the Large Hadron Collider.

DA field representation

• An analytic magnet model is needed (for example the LHC HGQ lead and return ends; by G. Sabbi, described in the Fermilab report TD-97-040)

• It is based on the Biot-Savart law for a piece of straight current wire in $3D$

• A naive implementation

$$\vec{B} = \frac{kI\left(\vec{l}\times\vec{r}_s\right)}{b^2-4ac}\cdot\left(\frac{2b}{\sqrt{a}}-\frac{2b}{\sqrt{a+b+c}}-\frac{4c}{\sqrt{a+b+c}}\right)$$

We used the shorthand notations $k=-\mu_0/(4\pi)$, $a=\left|\vec{r}_s\right|^2$, $b=2\vec{r}_s\cdot\vec{l}$, and $c=\left|\vec{l}\right|^2$.

• Differential Algebraic (DA) field computation consists of evaluation in DA of the above formula, and summation over all the contributions of the individual wires. This yields the values and all derivatives to any order of the fields.

Numerical pitfalls

Are due to cancellation:

$$\vec{B} = \frac{kI\left(\vec{l}\times\vec{r}_s\right)}{b^2-4ac}\cdot\left(\frac{2b}{\sqrt{a}}-\frac{2b}{\sqrt{a+b+c}}-\frac{4c}{\sqrt{a+b+c}}\right)$$

- If $b+c \ll a$ (short piece far from the reference point)
- If $b^2 \approx 4ac$ (wire (almost) parallel with the starting position vector)

Can be eliminated with the help of various simple tricks; numerically stable result is

$$\vec{B} = \frac{kI\left(\vec{l}\times\vec{r}_s\right)}{\left|\vec{r}_s\right|^2\left|\vec{r}_s+\vec{l}\right|\left(\left|\vec{r}_s+\vec{l}\right|+\left|\vec{r}_s\right|\right)}\cdot$$
$$\left[\left|\vec{r}_s\right|+\frac{\left|\vec{r}_s\right|^2\cos^2\theta+\left|\vec{r}_s+\vec{l}\right|^2}{\left|\vec{r}_s\right|\cos^2\theta+\left|\vec{l}\right|\cos\theta+\left|\vec{r}_s+\vec{l}\right|}\right]$$

Comparison between Mathematica and COSY Infinity

Mathematica

• Computation of derivatives in: one variable (x), one field component (B_y), one piece of wire

Order	$\frac{\partial^n B_y}{\partial x^n}(0,0,0)$	
n	Lines of Fortran	Eval. time $[\mu\,\mathrm{sec}]$
0	4	300
1	20	800
2	69	2400
3	188	7000
4	457	16000
5	1009	36000
6	2078	72900
7	4059	140900
8	7567	267000
9	13603	480000

• Mathematica-computed derivatives also have cancellation problems

COSY Infinity

- Computation of derivatives under the same conditions

Order	$\frac{\partial^n B_y}{\partial x^n}(0,0,0)$
n	Evaluation time in μsec up to order n
1	5
5	7
10	11
15	21
20	48
25	94

- It is the only feasible method in all 3 variables

Comparison between the two implementations of the Biot-Savart law

- Numerically unstable form

$$\left(\nabla \times \vec{B}\right)_s$$

Coefficient	Order	Exp. (x, s, y)
0.3368312531613524	2	2 0 0
0.1135639355887008E-06	2	1 1 0
0.2237449336917052E-08	2	0 2 0
0.6101442684425235E-09	2	1 0 1
0.1121592163033647E-06	2	0 1 1
-.3368312554111546	2	0 0 2
0.8482986704194671E-06	3	3 0 0
-1010523.362231316	3	2 1 0
0.4430214737283222E-06	3	1 2 0
-.3117654705420136E-04	3	0 3 0
-.3840225446083422E-05	3	2 0 1
0.9449170335074086E-08	3	1 1 1
0.6295506663533956E-05	3	0 2 1
-.2090122526610116E-05	3	1 0 2
1010523.362276393	3	0 1 2
0.5070636007076246E-05	3	0 0 3

- **Numerically stable form**

$$\left(\nabla \times \vec{B}\right)_s$$

Coefficient	Order	Exp. (x, s, y)
0.6297175136893429E-07	2	2 0 0
0.1135643710736822E-06	2	1 1 0
0.2254658681977162E-08	2	0 2 0
0.6101880112296945E-09	2	1 0 1
0.1121600312625759E-06	2	0 1 1
-.6522627415961324E-07	2	0 0 2
-.1035163990081857E-06	3	3 0 0
0.3603849257016734E-05	3	2 1 0
0.2310457119847342E-05	3	1 2 0
0.2364236389995611E-07	3	0 3 0
-.1977820552667708E-05	3	2 0 1
0.9434835135380232E-08	3	1 1 1
0.2256179271853398E-05	3	0 2 1
-.1999900781868291E-05	3	1 0 2
-.3674740044701252E-05	3	0 1 2
-.9278849333327333E-07	3	0 0 3

Multipole expansion

• Scalar potential: in the region of interest $\nabla \times \vec{B} = 0$, thus the fields are derivable from a magnetic scalar potential that satisfies the Laplace equation

The general solution is

$$V_B = \sum_{k,l=0}^{\infty} \left(b_{k,l}(s) \sin l\phi + a_{k,l}(s) \cos l\phi \right) r^k$$

• Recurrence relations among coefficients

$$b_{l+2n,l}(s) = \frac{b_{l,l}^{(2n)}(s)}{\Pi_{\nu=1}^{n} \left(l^2 - (l+2\nu)^2 \right)}$$

$$a_{l+2n,l}(s) = \frac{a_{l,l}^{(2n)}(s)}{\Pi_{\nu=1}^{n} \left(l^2 - (l+2\nu)^2 \right)}$$

• All other coefficients that cannot be obtained by these relations are zero.

• Thus, the free parameters that can be used to satisfy boundary conditions are the $b_{l,l}(s)$ and $a_{l,l}(s)$. They are called multipole strengths. The non-vanishing coefficients that are induced by their s-dependence are called pseudo-multipoles.

Iterative multipole extraction

• The field components B_x and B_y in the plane $y = 0$ contain all the information

$$B_x(x, y = 0, s) = \sum_{l=1}^{\infty} g_l(x, s) \cdot x^{l-1}$$

$$B_y(x, y = 0, s) = \sum_{l=1}^{\infty} f_l(x, s) \cdot x^{l-1}$$

where

$$g_l(x, s) = \sum_{n=0}^{\infty} (l + 2n) a_{l+2n,l}(s) x^{2n}$$

$$f_l(x, s) = \sum_{n=0}^{\infty} l b_{l+2n,l}(s) x^{2n}$$

• The multipole strengths can be extracted iteratively starting with the dipole component.

Enforcing Maxwell's equations

• Magnet models are often not closed: image currents, "leads", separate treatment of the two ends (in case of LHC HGQs lead and return ends)

• Closing the model is required by: Maxwell's equations, numerical stability of the multipole extraction algorithms

• There are two methods to accomplish it

  • Physically add wires in such a way that all the loose ends close at "infinity"

  • By imposing that the fictitious closing wires should minimize the modification of the original fields in a neighborhood of the optical axis

First method (Local Maxwellification)

• Physically close the model by adding new wires

• It is enough to compute the field components in two variables (x, s) and obtain from them the field in the whole space. This is done by transforming the Laplace equation for the potential into a fixed point problem.

$$V(x,y,s) = V(x,0,s) + \int_0^y \left.\frac{\partial V(x,\widetilde{y},s)}{\partial \widetilde{y}}\right|_{\widetilde{y}=0} d\bar{y}$$
$$- \int_0^y \int_0^{\bar{y}} \left(\frac{\partial^2 V(x,\widetilde{y},s)}{\partial x^2} + \frac{\partial^2 V(x,\widetilde{y},s)}{\partial s^2} \right) d\widetilde{y} d\bar{y}$$

The two initial conditions are

$$\int_0^y \left.\frac{\partial V(x,\widetilde{y},s)}{\partial \widetilde{y}}\right|_{\widetilde{y}=0} d\bar{y} = y \cdot B_y(x,0,s)$$
$$V(x,0,s) = \int_0^x B_x(\bar{x},0,0) d\bar{x} + \int_0^s B_s(x,0,\bar{s}) d\bar{s}$$

• Once we have the potential we can recompute the fields in the whole space by mere differentiation

This is a simple way to correct for small numerical errors

• Local Maxwellification can be done the same way in $3D$

Second method (Global Maxwellification)

• No physical closings, but compute the fields in all 3 variables.

• If $\vec{B}\,(x, y, s)$ is the result such that $\nabla \cdot \vec{B} = 0$ and $\nabla \times \vec{B} \neq 0$, there exists $\vec{R}\,(x, y, s)$, which stands for the fictitious closings, such that

$$\nabla \cdot \vec{R} = 0$$

$$\nabla \times \left(\vec{B} + \vec{R}\right) = 0$$

It follows that in general

$$R_\phi(r, \phi, s) =$$

$$\sum_{l=0}^{\infty} \left(f(r, s) \sin l\phi + g(r, s) \cos l\phi\right) r^{l-1}$$

$$-B_\phi(r, \phi, s)$$

We obtain a minimal R_ϕ if we choose the free parameters in f and g to cancel the corresponding terms in B_ϕ.

Implementation

• The practical implementation in COSY in actually an analytic s-dependent Fourier transform of B_ϕ or B_r

• Once we have the multipoles, the out of axis expansion is performed, the potential built up, and the new field components computed by differentiation

• The new field will satisfy Maxwell's equations exactly, and will alter the original field minimally; we call the method **global Maxwellification**

• In case the magnet model is already Maxwellian, the two methods are completely equivalent

• Using any of the methods it is trivial (and very accurate) to check whether Maxwell's equations are obeyed

Examples of LHC multipoles

• Multipole strengths (normal and skew) have been computed in both ends up to 28-poles

• Derivatives of multipoles have been computed up to order 12

Fringe field maps

• The equations of motion are integrated in COSY, yielding the map

• There are two methods on which computation of maps is based: a) Multipoles interpolated using Gaussian interpolation, or b) the potential is interpolated by a derivative preserving method

• The two methods are equivalent

Tracking the LHC

• Fully accurate fringe field maps incorporated

Conclusion: *Fringe fields introduce important nonlinear effects, but are not a limiting factor for the dynamic aperture.*

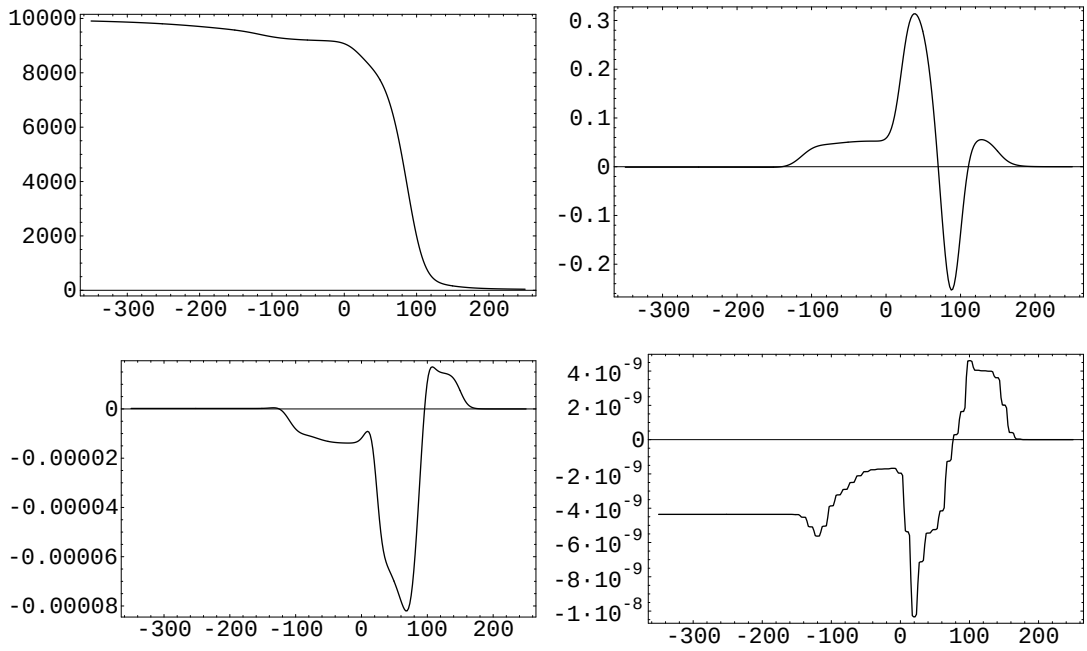

Figure 10: Normal multipoles of the lead end.

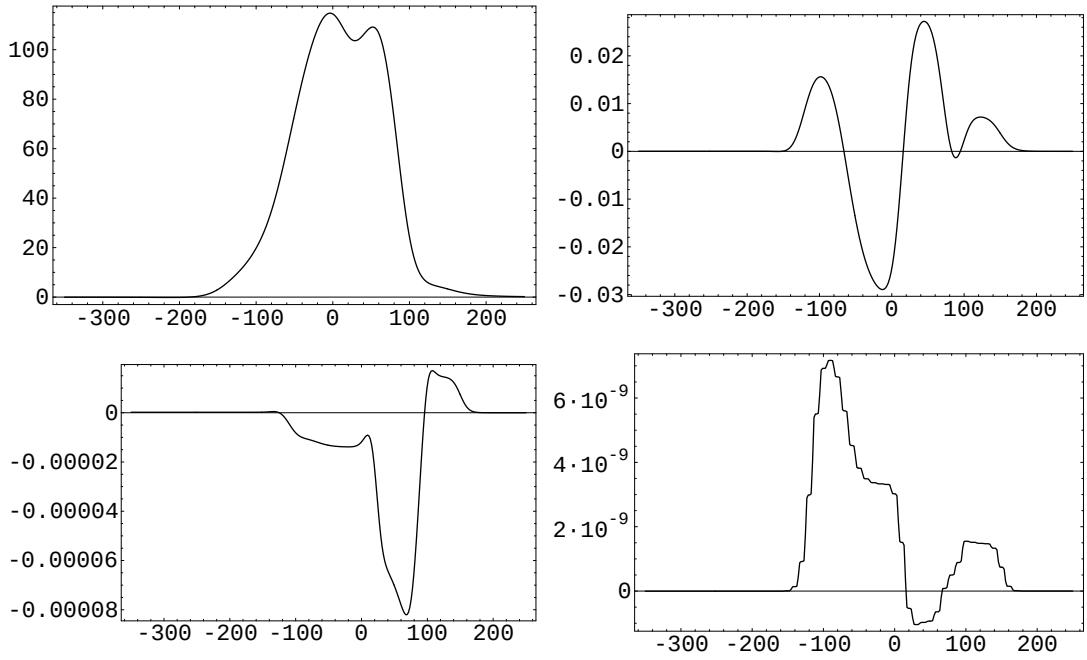

Figure 11: Skew multipoles of the lead end.

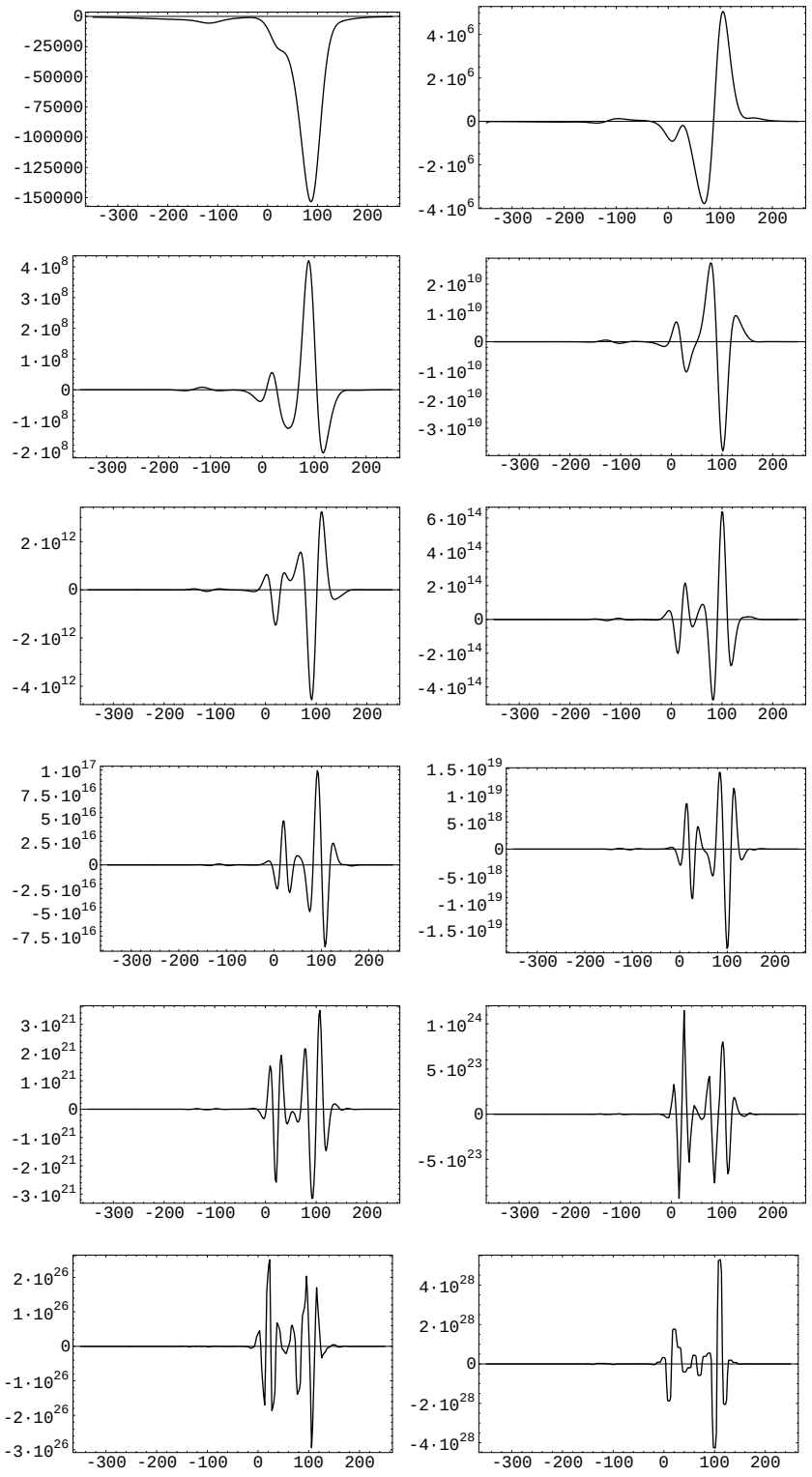

Figure 12: Axial derivatives (from first to 12th) of the normal quadrupole component.

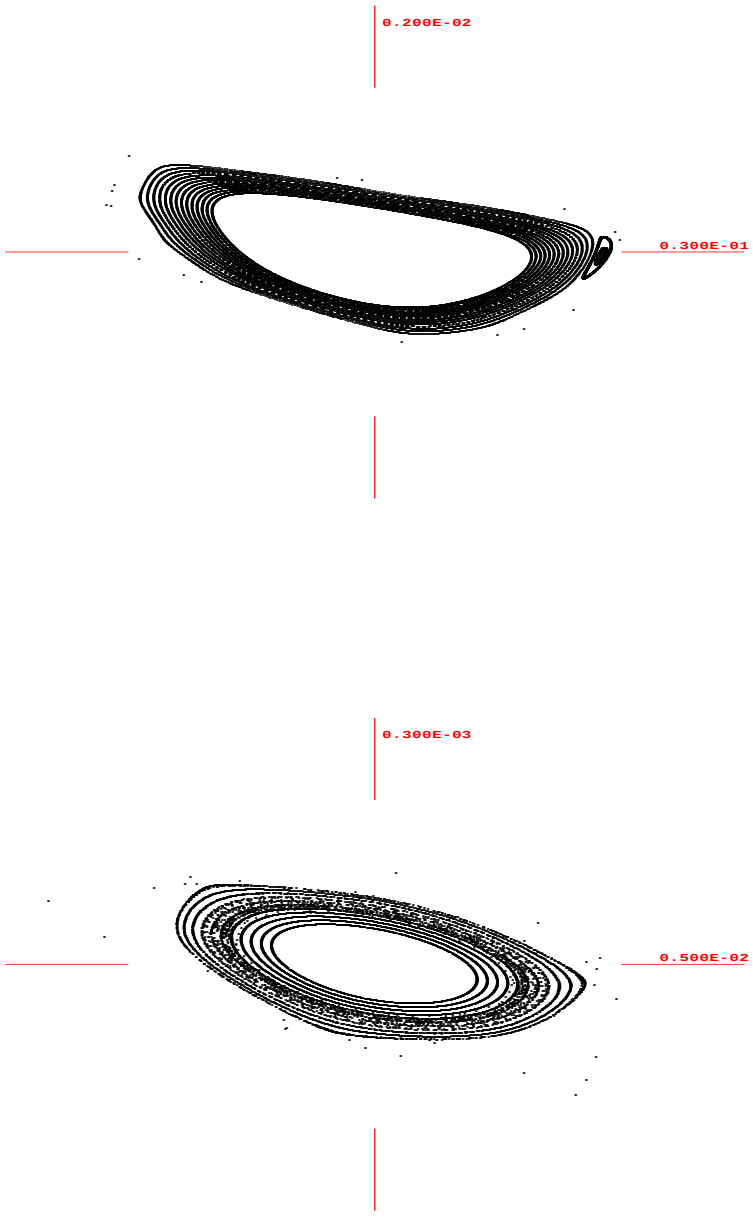

Figure 13: Symplectic tracking of the ideal lattice of the LHC without and with fringe fields.

Symplectification methods

Factorizations in "elementary" maps: Cremona, Solvable map, Integrable polynomial, Monomial, Kick, Jolt

Generating Functions: Conventional F_1 through F_4 (Goldstein's notation)

But there are more generating function types!

How many?

Which type of generating function is the best for symplectification purposes?

$\Rightarrow$ Develop the general theory of generating functions and formulate the condition for optimal symplectification

General theory of generating functions

Outline: Given a $2n$ dimensional symplectic map, look for scalar functions of $2n$ variables, defined on some space different from phase space, which give a representation of the symplectic map by some algorithm, through its partial derivatives.

Geometrization of the problem: The most general and transparent way to proceed is to transform the dynamical problem into a problem in symplectic geometry.

"Everything is Lagrangian": Symplectic maps, and scalar functions under certain conditions can be put into one-to-one correspondence with Lagrangian submanifolds of appropriate symplectic manifolds.

Identification of Lagrangian submanifolds: The most general diffeomorphism that sends the Lagrangian manifold determined by the symplectic map onto the Lagrangian submanifold determined by some scalar function extends to a local conformal symplectic map.

General theory of generating functions

If we denote the symplectic map by $\mathcal{M}$, we call F the generating function of type α of $\mathcal{M}$ if the following relationships hold:

$$(\nabla F)^T = \left(\alpha_1 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right) \circ \left(\alpha_2 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right)^{-1}$$

$$\mathcal{M} = \left(\alpha^1 \circ \begin{pmatrix} (\nabla F)^T \\ \mathcal{I} \end{pmatrix}\right) \circ \left(\alpha^2 \circ \begin{pmatrix} (\nabla F)^T \\ \mathcal{I} \end{pmatrix}\right)^{-1}$$

$$\text{where} \qquad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad , \quad \alpha^{-1} = \begin{pmatrix} \alpha^1 \\ \alpha^2 \end{pmatrix}.$$

$\mathcal{I}$ is the identity map, and α is a conformal symplectic map, meaning that it satisfies:

$$\left(\mathrm{Jac}\,(\alpha)\right)^T J_{4n}\, \mathrm{Jac}\,(\alpha) = \mu \tilde{J}_{4n},$$

for nonzero μ, and

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad , \quad \tilde{J}_{4n} = \begin{pmatrix} J_{2n} & 0_{2n} \\ 0_{2n} & -J_{2n} \end{pmatrix}.$$

Equivalence classes of generating functions

• Two generator types are equivalent if they produce exactly the same symplectified map

• The set of generating functions is degenerate from the symplectification point of view

• The pool of generators reduces to equivalence classes

For a given symplectic map $\mathcal{M}$ with linear part M, let β_1, β_2 be linear conformal symplectic maps

$$\mathrm{Jac}\,(\beta_{1,2}) = \begin{pmatrix} A_{1,2} & B_{1,2} \\ C_{1,2} & D_{1,2} \end{pmatrix},$$

such that $C_{1,2}M + D_{1,2}$ is invertible. Let

$$S(\beta_{1,2}) = -J\left(C_{1,2}M + D_{1,2}\right)^{-1}\left(C_{1,2}M - D_{1,2}\right).$$

Then, F_{β_1} is equivalent with F_{β_2} if and only if

$$S(\beta_1) = S(\beta_2).$$

• Observation:

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} -JM^{-1} & J \\ \frac{1}{2}\left(I + JS(\beta_{1,2})\right)M^{-1} & \frac{1}{2}\left(I - JS(\beta_{1,2})\right) \end{pmatrix}$$

produces an equivalent generator type.

Equivalence of symplectification procedures with/without the linear part

$$\mathcal{M} = M + H,$$

• We can distinguish three symplectification procedures: symplectify $\mathcal{M}_n$ directly (using some S), symplectify $\mathcal{M}_{L,n}$ (using some other $\bar{S}$) obtained from

$$\mathcal{M}_L = I + M^{-1} \circ H,$$

or symplectify $\mathcal{M}_{R,n}$ (using yet some other $\tilde{S}$) obtained from

$$\mathcal{M}_R = I + H \circ M^{-1}.$$

• Relation among the maps

$$\mathcal{M} = M \circ \mathcal{M}_L,$$
$$\mathcal{M} = \mathcal{M}_R \circ M.$$

• **Question**: Do these relations continue to hold for the symplectified versions of the corresponding maps?

• **Answer:** Yes, if and only if

$$S = \bar{S},$$
$$S = M^T \tilde{S} M.$$

Hofer's metric and optimal symplectification

- We need to define some "closeness" criterion
- Desirable properties of the metric:

  ◊ Symplectification should work well for any particle in a given Poincaré section

  ◊ Coordinate independence, meaning that what is ideal for $\mathcal{M} \Rightarrow$ ideal for $\mathcal{A} \circ \mathcal{M} \circ \mathcal{A}^{-1}$, for any symplectic $\mathcal{A}$

Solution: Hofer's metric

Definition: *For any two* $\varphi, \psi \in \mathrm{Ham}_c\left(\mathbb{R}^{2n}\right)$, *we define the distance between them as*

$$\rho\left(\varphi, \psi\right) = \inf_{\phi_0 = \varphi, \phi_1 = \psi} \int_0^1 \left(\max_{z \in \mathbb{R}^{2n}} H_t\left(z\right) - \min_{z \in \mathbb{R}^{2n}} H_t\left(z\right) \right) dt.$$

The infimum is taken over all smooth paths in $\mathrm{Ham}_c\left(\mathbb{R}^{2n}\right)$ *from* φ *to* ψ. *Here* $||H_t|| = \max_{z \in \mathbb{R}^{2n}} H_t\left(z\right) - \min_{z \in \mathbb{R}^{2n}} H_t\left(z\right)$ *is called the oscillation norm.*

Hofer's metric

Hofer's metric is an essentially unique

- ◊ Intrinsic
- ◊ Bi-invariant
- ◊ Finsler

metric for Hamiltonian symplectic maps.

Definition: A symplectified map is called optimal, if the distance in Hofer's metric between the exact map and the symplectified map is minimized over the set of all symplectified maps.

Theorem: *There exists a neighborhood* $\mathcal{E}$ *of any* $\mathcal{M} \in \mathrm{Ham}_c\left(\mathbb{R}^{2n}\right)$, *and a neighborhood* $\mathcal{Z}$ *of* 0 *in* $C_c^\infty\left(\mathbb{R}^{2n}\right)$ *such that the map*

$$\Phi_\alpha : \mathcal{Z} \to \mathcal{E} \qquad \Phi_\alpha\left(F\right) = \mathcal{M}$$

is isometric for any α, *as long as the corresponding generating function type exists for both symplectic maps. That is, for every* $F, G \in \mathcal{Z}$,

$$||F - G|| = |\mu| \cdot \rho\left(\Phi_\alpha\left(F\right), \Phi_\alpha\left(G\right)\right).$$

Optimal symplectification

- In practice it is straightforward to compute F_n
- Applying Hofer's metric to symplectic tracking of accelerators, we obtain

$$\rho\left(\Phi_\alpha\left(F\right), \Phi_\alpha\left(F_n\right)\right) = ||F - F_n||.$$

Conclusion: *for compactly supported Hamiltonian symplectomorphisms, the optimal symplectification using the mixed variable generating function method is achieved by the order $n+1$ truncated generating function that has the* ***smallest oscillation norm*** *of the terms neglected, above order $n+1$, and therefore minimizes the right hand side of the above equation.*

Choosing the best generating function

- Minimization of $||F - F_n|| \Leftrightarrow$ Minimization of $||F||$
- Estimate F from

$$F(w) = \int_0^w \alpha_1(\hat{z}, z) \cdot dw,$$

where

$$\hat{z} = \mathcal{M}(z),$$
$$w = \alpha_2(\hat{z}, z).$$

It follows that

$$\| F \| \leq \| \alpha_1 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} \| \cdot \| \alpha_2 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} \|.$$

- Moreover,

$$\alpha_1(\hat{z}, z) = 0 + \mathcal{O}(z^2),$$
$$\alpha_2(\hat{z}, z) = I \cdot z + \frac{1}{2}(I + JS) \cdot \mathcal{O}(z^2).$$

EXPO

In general (statistically, or on average), optimal symplectification is achieved by the class $[S]$ obeying

$$S = 0,$$

corresponding to the generating function type associated with

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} -JM^{-1} & J \\ \frac{1}{2}M^{-1} & \frac{1}{2}I \end{pmatrix}.$$

We call it the **EX**tended **PO**incaré (**EXPO**) type.

Examples

1. A symplectic map generated from a random Hamiltonian.

2. An anharmonic oscillator:

$$H = \frac{1}{2}\left(p^2 + q^2\right) - \frac{1}{4}q^4.$$

3. An exactly symplectic quadratic map:

$$\mathcal{M} = \mathcal{N} \circ \mathcal{L},$$

where

$$\mathcal{L} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

with $\theta = \frac{\pi}{3}$, and

$$\mathcal{N}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q - 3\left(q+p\right)^2 \\ p + 3\left(q+p\right)^2 \end{pmatrix}.$$

4. A lattice of the proposed Neutrino Factory, and the Proton Driver.

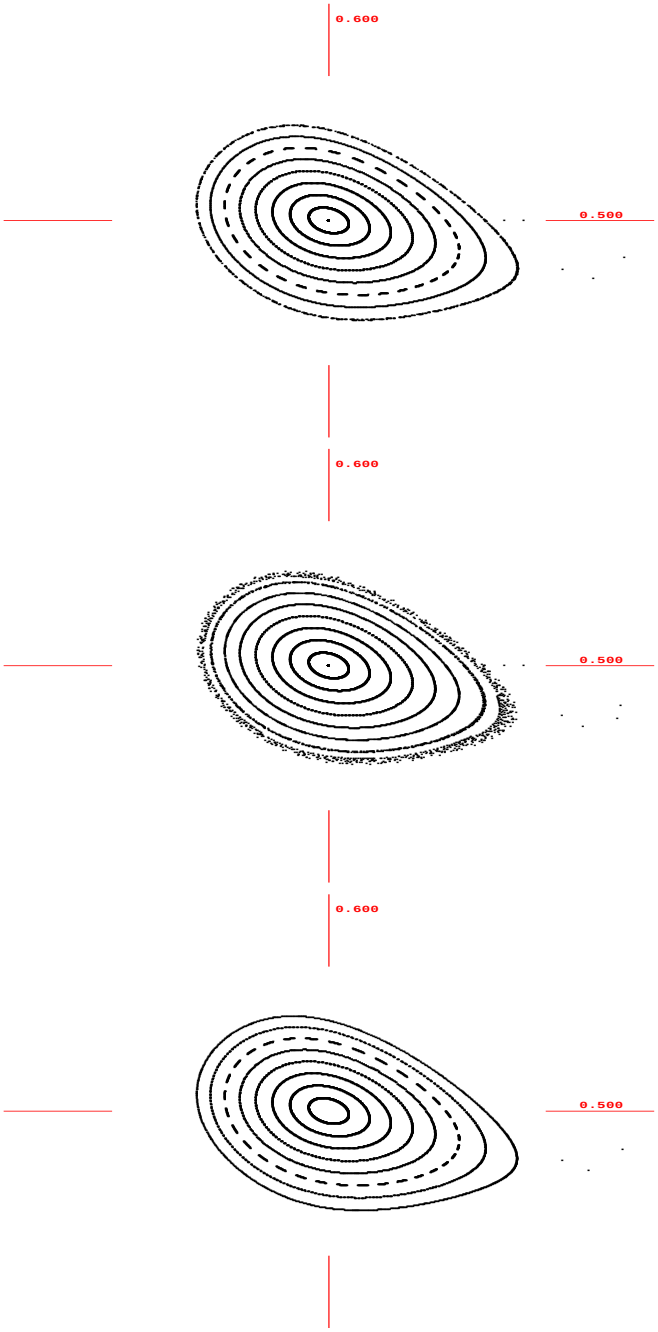

Figure 14: Non-symplectic tracking with Taylor maps of orders 19 and 11, and order 11 symplectic tracking with EXPO for a $2D$ symplectic map obtained from a random Hamiltonian with coefficients in $[-1, 1]$.

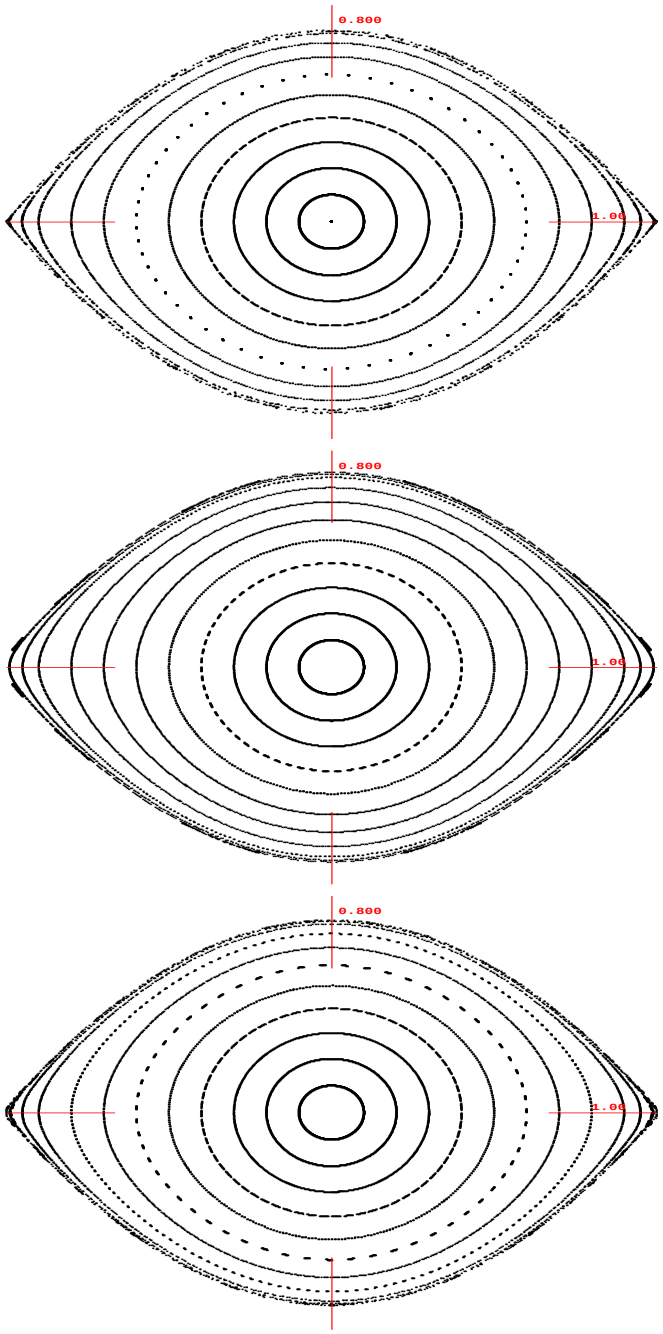

Figure 15: Non-symplectic tracking with Taylor map of order 19, and order 7 and 11 symplectic tracking with EXPO, for an anharmonic oscillator.

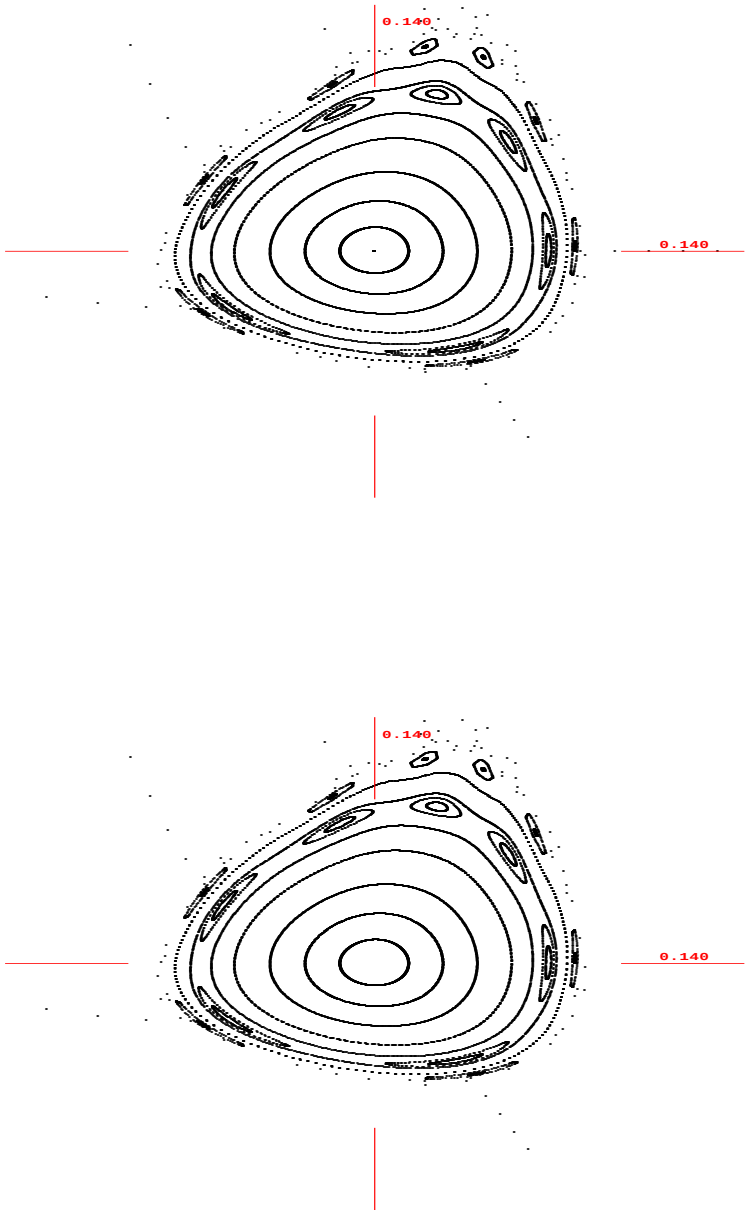

Figure 16: Non-symplectic tracking with Taylor map of order 3, and order 3 symplectic tracking with EXPO, for an exactly symplectic quadratic map.

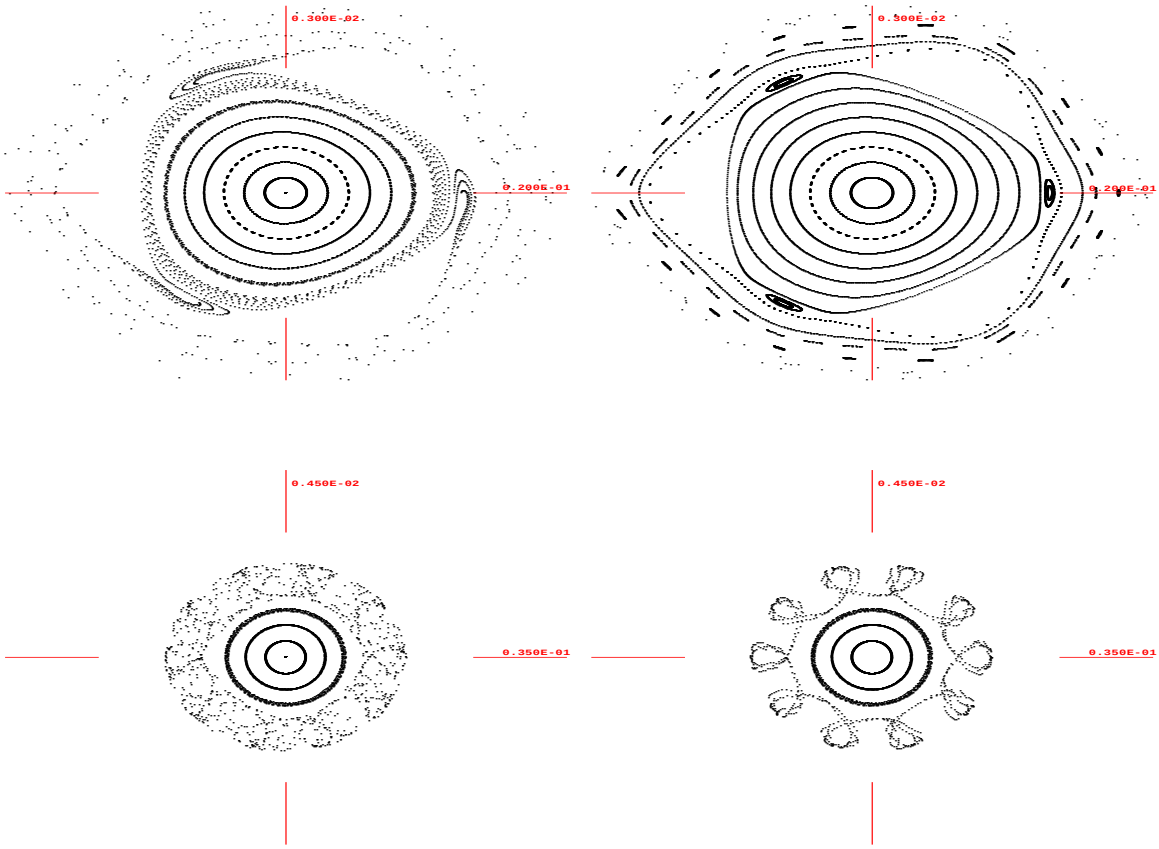

Figure 17: Non-symplectic tracking with Taylor map of order 8, and order 8 symplectic tracking with EXPO, for two realizations of a lattice of the proposed Neutrino Factory.

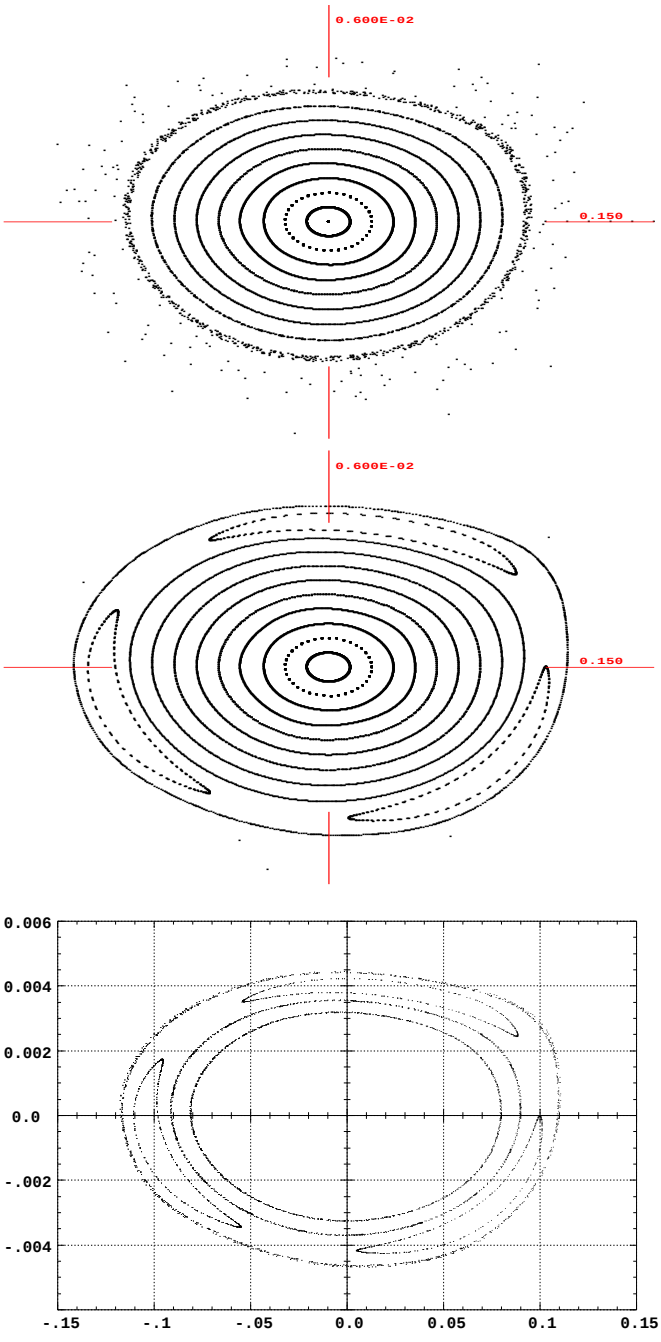

Figure 18: 1000 turn tracking of the FNAL Proton Driver with the 15th order Taylor map, the corresponding $S = 0$ symplectified map, and the element-by-element numerical integration.

Summary

- Consideration of fringe field effects is always indicated
- It is possible to compute "exact" fringe field maps
- Tracking with fringe fields requires the same effort as without fringe fields
- There are many more generating function types than commonly known
- Optimal symplectification can be formulated using techniques of symplectic geometry and topology
- In general, EXPO is the optimal generating function type